

WHAT IS CLAIMED IS:

1. A method for generating a data stream according to a binary format of a tag-based description language, comprising:

tokenizing tag names into numeric tokens.

5

2. A method according to claim 1, further comprising:

tokenizing attribute names into numeric tokens.

3. A method according to claim 1, wherein said numeric tokens for tag names are variable sized.

10

4. A method according to claim 2, wherein said numeric tokens for attribute names are variable sized.

5. A method according to claim 2, wherein said tokenizing of attributes enables values natively stored as binary data types to be inserted into the data stream.

15

6. A method according to claim 1, wherein said tokenizing of tag names includes inserting a name definition construct into the data stream the first time a tag name is encountered for purposes of recreating the tag names by a device that receives the data stream.

20

7. A method according to claim 1, wherein the tag-based description language is extensible markup language (XML).

25

8. A method according to claim 2, wherein the tokenizing of the tag and attribute names decreases the time elapsed parsing the data stream by a device that receives the data stream, the time being decreased relative to the parsing of a corresponding text-based format of the tag-based description language.

30

9. A method according to claim 2, wherein the tokenizing of the tag and attribute names decreases overhead incident to formatting data for representation according to the tag-based description language.

5 10. A method according to claim 2, wherein the tokenizing of the tag and attribute names decreases the size of the resulting data file formatted according to the tag-based description language.

11. A computer readable medium bearing computer executable instructions for carrying
10 out the method of claim 1.

12. A computer readable medium bearing computer executable instructions for carrying
out the method of receiving a well-formed document in a text format of a tag-based
description language and converting the document to a binary format via tokenization of the
tag and attribute names into numeric tokens.
15

13. A computer readable medium according to claim 12, wherein said tokenization of
attributes enables values natively stored as binary data types to be inserted into the data
stream.
20

14. A computer readable medium according to claim 12, wherein said tokenization of tag
names includes inserting a name definition construct into the data stream the first time a tag
name is encountered for purposes of recreating the tag names by a device that receives the
data stream.
25

15. A computer readable medium according to claim 12, said receiving includes receiving
a document formatted according to a text format of XML.

16. A computer readable medium bearing computer executable instructions for carrying
30 out the method of assembling data into a document according to a binary format by

tokenizing the tag and attribute names into variable sized numeric tokens.

17. A computer readable medium according to claim 16, wherein said tokenizing of attributes enables values natively stored as binary data types to be inserted into the data stream.

18. A computer readable medium according to claim 16, wherein said tokenizing of tag names includes inserting a name definition construct into the data stream the first time a tag name is encountered for purposes of recreating the tag names by a device that receives the data stream.

19. A computer readable medium according to claim 16, said receiving includes receiving a document formatted according to a text format of XML.

20. A computer readable medium bearing computer executable instructions for carrying out the method of receiving a document formatted according to a binary format of a tag-based description language and directly parsing the data in the document for use by another element in a computer system.

21. A computer readable medium according to claim 20, wherein before said parsing, said method includes converting the document to a text format of the tag-based description language.

22. A computer readable medium according to claim 20, wherein said receiving includes receiving a document formatted according to a binary format of XML.

23. A computing device, comprising:

means for receiving a well-formed document in a text format of a tag-based description language;

means for converting the document to a binary format via tokenization of the tag and

attribute names into variable sized numeric tokens; and

means for converting the document back to the text format without a loss of fidelity.

24. A computing device according to claim 23, wherein said tokenization of attributes
5 enables values natively stored as binary data types to be inserted into the data stream.

25. A computing device according to claim 23, wherein said tokenization of tag names
includes inserting a name definition construct into the data stream the first time a tag name is
encountered for purposes of recreating the tag names by a device that receives the data
10 stream.

26. A computing device according to claim 23, said tag-based description language is
XML.

27. In a system in which a transmitting device transmits in a streaming fashion data
formatted according to a tag-based description language, a method for generating a data
stream according to a binary format of the tag-based description language, comprising:
15

for each unique tag name, at the first time a tag name of the data is encountered,
tokenizing the tag name into a numeric token and transmitting the token and the text
associated with the tag name; and
20

at any time subsequent to the first time that the tag name of the data is encountered,
transmitting the numeric token without the text.

28. A method according to claim 27, further comprising:
25 tokenizing attribute names into numeric tokens.

29. A method according to claim 27, wherein said data is transmitted incrementally, and
whereby a receiving device parses said data as it is incrementally received by the receiving
device.
30

30. A method according to claim 27, wherein said numeric tokens for tag names are variable sized.

31. A method according to claim 28, wherein said numeric tokens for attribute names are variable sized.

32. A method according to claim 28, wherein said tokenizing of attributes enables values natively stored as binary data types to be inserted into the data stream.

33. A method according to claim 27, wherein the tag-based description language is extensible markup language (XML).

34. A method according to claim 28, wherein the tokenizing of the tag and attribute names decreases the time elapsed parsing the data stream by a device that receives the data stream, the time being decreased relative to the parsing of a corresponding text-based format of the tag-based description language.

35. A method according to claim 28, wherein the tokenizing of the tag and attribute names decreases overhead incident to formatting data for representation according to the tag-based description language.

36. A method according to claim 28, wherein the tokenizing of the tag and attribute names decreases the size of the resulting data file formatted according to the tag-based description language.

37. A computer readable medium bearing computer executable instructions for carrying out the method of claim 27.

38. A method for generating a data stream according to an XML binary format, comprising:

tokenizing tag names and attribute names into variable sized numeric tokens,
wherein said tokenizing of attributes enables values natively stored as binary data
types to be inserted into the data stream, wherein said tokenizing of tag names includes
inserting a name definition construct into the data stream the first time a tag name is
5 encountered for purposes of recreating the tag names by a device that receives the data
stream, thereby decreasing parsing time.